

CLAIMS

What is claimed is:

1. A system for identification and characterization of gene expression in one or more samples, comprised of:

- (a) providing one or more samples comprising one or more mRNA molecules;
- (b) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identimer also comprising a detectable marker at its 5' end;
- (c) contacting said mRNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) reverse transcribing the mRNA to produce a first strand cDNA that includes the identimer;
- (e) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (f) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;

- (g) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments; and
- (h) amplifying the one or more ligated cleavage fragments using the identimer to produce one or more amplified fragments comprising sequences complementary to a 3' end of the mRNA.

2. The system of Claim 1, further comprising the identification and characterization of the cleavage fragments according to the presence of the marker, the sequences corresponding to the VN_x nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

3. The system of Claim 2, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database contacting sequence and size characteristics of RNAs associated with known genes.

4. The system of Claim 3 whereby said comparison is conducted by means of software operated on a microprocessor.

5. A system for identification and characterization of gene expression in one or more samples, comprised of:

- (a) providing one or more samples comprising one or more mRNA molecules;
- (b) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VNNN, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C,

G, or T, said identimer also comprising a detectable marker at its 5' end;

- (c) contacting said mRNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VNNN portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) reverse transcribing the mRNA to produce a first strand cDNA that includes the identimer;
- (e) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (f) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (g) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments; and
- (h) amplifying the one or more ligated cleavage fragments using the identimer to produce one or more amplified fragments comprising sequence complementary to a 3' end of the mRNA.

6. The system of Claim 5, further comprising the identification and characterization of the cleavage fragments according to the presence of the marker, the sequences corresponding to the VNNN nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

7. The system of Claim 6, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database containing sequence and size characteristics of RNAs associated with known genes.

8. The system of Claim 7 whereby said comparison is conducted by means of software operated on a microprocessor.

9. A system for identification and characterization of gene expression in two or more samples, comprised of:

- (a) providing a first sample comprising one or more mRNA molecules;
- (b) providing an identifier comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identifier also comprising a first detectable marker at its 5' end;
- (c) contacting said mRNA in the first sample with the identifier comprising said first detectable marker such that the polyT portion of the identifier hybridizes to the polyA tail of the mRNA and the VN_x portion of the identifier hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) providing a second sample comprising one or more mRNA molecules;

- (e) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identimer also comprising a second detectable marker at its 5' end;
- (f) contacting said mRNA in the second sample with the identimer comprising said second detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (g) reverse transcribing the mRNA in each sample to produce a first strand cDNA that includes the respective identimer;
- (h) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (i) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (j) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments; and
- (k) amplifying the one or more ligated cleavage fragments using the respective identimer to produce one or more amplified fragments comprising sequence complementary to a 3' end of the mRNA.

10. The system of Claim 9, further comprising the identification and characterization of the cleavage fragments according to the presence of the marker, the sequences corresponding to the VN_x nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

11. The system of Claim 10, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database containing sequence and size characteristics of RNAs associated with known genes.

12. The system of Claim 11 whereby said comparison is conducted by means of software operated on a microprocessor.

13. A system for identification and characterization of gene expression in two or more samples, comprised of:

- (a) providing a first sample comprising one or more mRNA molecules;
- (b) providing an identifier comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_{nn}, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C, G, or T, said identifier also comprising a first detectable marker at its 5' end;
- (c) contacting said mRNA in the first sample with the identifier comprising said first detectable marker such that the polyT portion of the identifier hybridizes to the polyA tail of the mRNA and the VN_{nn} portion of the identifier hybridizes with portions of the mRNA immediately upstream of the polyA tail;

- (d) providing a second sample comprising one or more mRNA molecules;
- (e) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VNNN, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C, G, or T, said identimer also comprising a second detectable marker at its 5' end;
- (f) contacting said mRNA in the second sample with the identimer comprising said second detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (g) reverse transcribing the mRNA in each sample to produce a first strand cDNA that includes the respective identimer;
- (h) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (i) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (j) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments; and
- (k) amplifying the one or more ligated cleavage fragments using the respective identimer to produce one or more amplified fragments comprising sequence complementary to a 3' end of the mRNA.

14. The system of Claim 13, further comprising the identification and characterization of the cleavage fragments according to the presence of the marker, the sequences corresponding to the VNNN nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

15. The system of Claim 14, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database containing sequence and size characteristics of RNAs associated with known genes.

16. The system of Claim 15 whereby said comparison is conducted by means of software operated on a microprocessor.

17. A system for identification and characterization of gene expression in one or more samples, comprised of:

- (a) providing one or more samples comprising one or more mRNA molecules;
- (b) providing an identifier comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identifier also comprising a detectable marker at its 5' end;

- (c) contacting said mRNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) reverse transcribing the mRNA to produce a first strand cDNA that includes the identimer;
- (e) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (f) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (g) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments;
- (h) amplifying the one or more ligated cleavage fragments by means of in vitro transcription using one or more RNA polymerases to produce in vitro transcribed RNA;
- (i) contacting said in vitro transcribed RNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the in vitro transcribed RNA and the VN_x portion of the identimer hybridizes with portions of the in vitro transcribed RNA immediately upstream of the polyA tail; and
- (j) reverse transcribing the in vitro transcribed RNA to produce a first strand cDNA that includes the identimer.

18. The system of Claim 17, further comprising the identification and characterization of the cleavage fragments represented by the first strand cDNA according to the presence of the marker, the sequences corresponding to the VN_x nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

19. The system of Claim 18, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database contacting sequence and size characteristics of RNAs associated with known genes.

20. The system of Claim 19 whereby said comparison is conducted by means of software operated on a microprocessor.

21. A system for identification and characterization of gene expression in one or more samples, comprised of:

- (a) providing one or more samples comprising one or more mRNA molecules;
- (b) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VNNN, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C, G, or T, said identimer also comprising a detectable marker at its 5' end;

- (c) contacting said mRNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VNNN portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) reverse transcribing the mRNA to produce a first strand cDNA that includes the identimer;
- (e) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (f) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (g) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments;
- (h) amplifying the one or more ligated cleavage fragments by means of in vitro transcription using one or more RNA polymerases to produce in vitro transcribed RNA;
- (i) contacting said in vitro transcribed RNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the in vitro transcribed RNA and the VN_x portion of the identimer hybridizes with portions of the in vitro transcribed RNA immediately upstream of the polyA tail; and
- (j) reverse transcribing the in vitro transcribed RNA to produce a first strand cDNA that includes the identimer.

22. The system of Claim 21, further comprising the identification and characterization of the cleavage fragments represented by the first strand cDNA according to the presence of the marker, the sequences corresponding to the VNNN nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

23. The system of Claim 22, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database contacting sequence and size characteristics of RNAs associated with known genes.

24. The system of Claim 23 whereby said comparison is conducted by means of software operated on a microprocessor.

25. A system for identification and characterization of gene expression in two or more samples, comprised of:

- (a) providing a first sample comprising one or more mRNA molecules;
- (b) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identimer also comprising a first detectable marker at its 5' end;

- (c) contacting said mRNA in the first sample with the identimer comprising said first detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) providing a second sample comprising one or more mRNA molecules;
- (e) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identimer also comprising a second detectable marker at its 5' end;
- (f) contacting said mRNA in the second sample with the identimer comprising said second detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (g) reverse transcribing the mRNA in each sample to produce a first strand cDNA that includes the respective identimer;
- (h) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (i) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;

- (j) ligating an adaptamer comprising an RNA polymerase promoter site to one or more of said cleavage fragments;
- (k) amplifying the one or more ligated cleavage fragments by means of in vitro transcription using one or more RNA polymerases to produce in vitro transcribed RNA;
- (l) contacting said in vitro transcribed RNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the in vitro transcribed RNA and the VNx portion of the identimer hybridizes with portions of the in vitro transcribed RNA immediately upstream of the polyA tail; and
- (m) reverse transcribing the in vitro transcribed RNA to produce a first strand cDNA that includes the identimer.

26. The system of Claim 25, further comprising the identification and characterization of the cleavage fragments represented by the first strand cDNA according to the presence of the marker, the sequences corresponding to the VNx nucleotide sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

27. The system of Claim 26, further comprising the identification of any gene associated with the cleavage fragments by comparing the sequence and size characteristics of the cleavage fragment with a database containing sequence and size characteristics of RNAs associated with known genes.

28. The system of Claim 27 whereby said comparison is conducted by means of software operated on a microprocessor.

29. A system for identification and characterization of gene expression in two or more samples, comprised of:

- (a) providing a first sample comprising one or more mRNA molecules;
- (b) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VNNN, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C, G, or T, said identimer also comprising a first detectable marker at its 5' end;
- (c) contacting said mRNA in the first sample with the identimer comprising said first detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VN_x portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (d) providing a second sample comprising one or more mRNA molecules;
- (e) providing an identimer comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VNNN, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, and each N equals a nucleotide A, C, G, or T, said identimer also comprising a second detectable marker at its 5' end;

- (f) contacting said mRNA in the second sample with the identimer comprising said second detectable marker such that the polyT portion of the identimer hybridizes to the polyA tail of the mRNA and the VNNN portion of the identimer hybridizes with portions of the mRNA immediately upstream of the polyA tail;
- (g) reverse transcribing the mRNA in each sample to produce a first strand cDNA that includes the respective identimer;
- (h) synthesizing a second DNA strand complementary to the first strand cDNA to form a duplex;
- (i) cleaving the duplex with at least one sequence-specific cleaving agent to provide one or more duplex cleavage fragments;
- (j) ligating an adaptor comprising an RNA polymerase promoter site to one or more of said cleavage fragments;
- (k) amplifying the one or more ligated cleavage fragments by means of in vitro transcription using one or more RNA polymerases to produce in vitro transcribed RNA;
- (l) contacting said in vitro transcribed RNA with said identimer such that the polyT portion of the identimer hybridizes to the polyA tail of the in vitro transcribed RNA and the VN_x portion of the identimer hybridizes with portions of the in vitro transcribed RNA immediately upstream of the polyA tail; and
- (m) reverse transcribing the in vitro transcribed RNA to produce a first strand cDNA that includes the identimer.

30. The system of Claim 29, further comprising the identification and characterization of the cleavage fragments represented by the first strand cDNA according to the presence of the marker, the sequences corresponding to the VNNN base pair sequence and the sequence associated with the sequence-specific cleaving agent, and the size of the fragment.

31. The system of Claim 30, further comprising the identification of any gene associated with the cleavage fragment by comparing the sequence and size characteristics of the cleavage fragments with a database contacting sequence and size characteristics of RNAs associated with known genes.

32. The system of Claim 31 whereby said comparison is conducted by means of software operated on a microprocessor.

33. A kit comprising:

- (a) one or more identimers comprising an oligo-dT primer of sequence, from 5' to 3' end, of T_n-VN_x, where n is an integer 8 or greater but not more than 50 representing the number of T's, V equals a nucleotide A, C, or G but not T, each N equals a nucleotide A, C, G, or T, and x is an integer 3 or greater but not more than 10 representing the number of N nucleotides, said identimer also comprising a detectable marker at its 5' end; and
- (b) one or more sequence-specific cleaving agents.